

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2002, 03:17:36 ; Search time 39 Seconds
(without alignments)

202.064 Million cell updates/sec

Title: us-09-895-298a-83
Perfect score: 1002
Sequence: 1 MMNFQPPSKAMRASQMMTF.....HDCSLDLRSRSVQEGNRA 130

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99.5	9.9	1203	YT41_CAEEL	Q11069 caenorhabdi
2	84	8.4	399	YXAM_BACSU	P42112 bacillus su
3	79.5	7.9	149	Y38A_MYCSE	O92b71 mycoplasma
4	79	7.9	239	PNUC_SALTY	P24520 salmonella
5	77	7.7	342	SPAS_SHIFL	P40707 shigella fl
6	77	7.7	564	YEDQ_EC057	Q88b92 escherichia
7	76.5	7.6	220	T2N5_NOSS7	P35677 nostoc sp.
8	76	7.6	564	YEDQ_EC057	P76330 escherichia
9	75.5	7.5	1189	ALAB_ARATH	O91k90 arabidopsis
10	75.5	7.5	2410	MOK1_SCHPO	O9nsk8 schizosach
11	75	7.5	422	PAN1_HUMAN	O96rd7 homo sapien
12	75	7.5	784	TLR2_BOVIN	O951a9 bos taurus
13	75	7.5	784	TLR2_HUMAN	O60603 homo sapien
14	74.5	7.4	148	DSBI_BACSU	O64037 bacillus su
15	74.5	7.4	385	Y464_MYCSE	P47702 mycoplasma
16	74.5	7.4	385	Y464_MYCSE	P75112 mycoplasma
17	74.5	7.4	401	OPRM_BOVIN	P79350 bos taurus
18	74.5	7.4	433	YBL4_YEAST	P38211 saccharomyc
19	73.5	7.3	159	LSPA_STACA	O59835 staphylococ
20	73.5	7.3	616	YGO2_YEAST	P30777 saccharomyc
21	73	7.3	284	COX3_LEITA	P14546 leishmania
22	73	7.3	496	GAB3_DROME	Q08832 drosophila
23	73	7.3	516	CALR_RAT	P32214 rattus norv
24	72.5	7.2	568	PTLB_STRMU	P50976 streptococc
25	72.5	7.2	830	YOG2_YEAST	P40367 saccharomyc
26	72	7.2	362	AOX_NEUCR	O01355 neurospora
27	72	7.2	444	CYB_RHOSH	O02761 rhodospira
28	71.5	7.1	500	CSL1_BACSU	P45860 bacillus su
29	71.5	7.1	971	Y277_MYCPN	P75387 mycoplasma
30	71.5	7.1	1064	CY44_RAT	P26770 rattus norv
31	71.5	7.1	1087	AKA9_RABIT	O28628 oryctolagus
32	71	7.1	189	Y78B_METJA	P81231 methanococc
33	71	7.1	239	PNUC_EC057	P31215 escherichia

34	70.5	7.0	208	1	ATPF_MYCSE	P47643 mycoplasma
35	70.5	7.0	390	1	EMRA_HAEIN	P44928 haemophilus
36	70.5	7.0	398	1	OPRM_MOUSE	P42866 mus musculus
37	70.5	7.0	398	1	OPRM_RAT	P35353 rattus norv
38	70.5	7.0	400	1	OPRM_HUMAN	P35372 homo sapien
39	70.5	7.0	689	1	DNLJ_RICPR	Q92ck9 rickettsia
40	70	7.0	226	1	ATP6_HELPY	Q92115 helicobacte
41	70	7.0	226	1	ATP6_HELPY	P56085 helicobacte
42	70	7.0	397	1	YMO4_CAEEL	O10907 caenorhabdi
43	70	7.0	510	1	EMRB_HAEIN	P44927 haemophilus
44	70	7.0	578	1	YGO2_EC057	P76007 escherichia
45	70	7.0	1342	1	ERB3_HUMAN	P21860 homo sapien

ALIGNMENTS

RESULT 1									
ID	YT41_CAEEL	STANDARD:	PRT:	1203 AA.					
AC	Q11069;	01-NOV-1997 (Rel. 35, Created)							
DT	16-OCT-2001 (Rel. 40, Last sequence update)								
DT	16-OCT-2001 (Rel. 40, Last annotation update)								
DE	Hypothetical 136.1 kDa protein B0416.1 in chromosome X.								
GN	B0416.1.								
OS	Caenorhabditis elegans.								
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;								
OC	Rhabditidae; Peloderinae; Caenorhabditis.								
OX	NCBI_Taxid=6239;								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=Bristol N2;								
RA	Favello T.;								
RL	Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.								
RP	REVISIONS.								
RA	Waterston R.;								
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/).								
CC	or send an email to license@isb-sib.ch .								
DR	EMBL: U23516; AAC38884.1; "								
DR	WormPep; B0416.1; CE19663.								
KW	Hypothetical protein.								
SQ	SEQUENCE 1203 AA; 136059 MW; 1B9686F827CC206 CRC64;								
Query Match 9.9%; Score 99.5; DB 1; Length 1203;									
Best Local Similarity 27.7%; Pred. NO.0.26;									
Matches 39; Conservative 23; Mismatches 50; Indels 29; Gaps 8;									
QY	6	PSKAMRASQMMTFIFLLFPSTGVLCTAI--TIWRLKPSADCGPRLPLFHSIY	63						
DB	701	PASQIFRASRSNFFALLILFLF--LCTLPVGVIAVKTPSKGCGPF-GNOSFEYSVI	756						
QY	64	SWIDPLSTRPGVILWVYIRNLIGSVHFF-----ITLVLVITYIYKQIT----	112						
DB	757	T--DVLHEN-----LDKTLVNGIKYSLSPGIIPVLVLSIVFLIAVTLGSQLAN	806						
QY	113	RKIMIRLLHEQIINEGDKMF	133						
DB	807	QDLSFQIMVER--TEKKKIF	825						
RESULT 2									
ID	YXAM_BACSU	STANDARD:	PRT:	399 AA.					

AC 042112;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein yxam.
 GN yxam OR S14MR
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 ON NCBI_TaxID=1423;
 RX MEDLINE=9603926; PubMed=7584049;
 RT Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
 RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
 RT genome between the gut and lol operons.";
 RL DNA Res. 2:61-69(1995).
 RN [2]
 RP REVISIONS.
 RA Fujita Y., Shibayama T., Ishio I., Aoyama D., Yoshida K.-I.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A.G., Borchert S.,
 RA Borstis R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capiano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denicot F., Devine K.M., Dusterhoft A., Enrich S.D., Emmerson P.T.,
 RA Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaeser P., Goffeau A., Golligly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Mestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Pario V., Pohl T.M., Portetle D., Portwolk S., Prescott A.M.,
 RA Pirescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandendol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AB005554; BAA21592.1; -
 DR EMBL: Z99124; CAB16029.1; -
 DR Subtilisin; BG1115; yxam.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 7 27
 FT TRANSMEM 33 53
 FT TRANSMEM 79 99
 FT POTENTIAL.
 FT POTENTIAL.

FT TRANSMEM 131 151 POTENTIAL.
 FT TRANSMEM 153 173 POTENTIAL.
 FT TRANSMEM 208 228 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 296 316 POTENTIAL.
 FT TRANSMEM 335 355 POTENTIAL.
 FT TRANSMEM 357 377 POTENTIAL.
 SQ SEQUENCE 399 AA; 44746 MW; 94AF8E563222E72 CRC64;
 Query Match 8.4%; Score 84; DB 1; Length 399;
 Best Local Similarity 19.9%; Pred. No. 2;
 Matches 35; Conservative 31; Mismatches 48; Indels 62; Gaps 8;
 QY 4 FQPPSKAM-----RASQMTFFFLI---FPPSF-TGVLCYTAITWRKPSADGGR 53
 DB 50 FEVPSGVFADRYGKIKISIASFFSILTWAFPPDISAICILAMITWALSDLSGFFE 109
 QY 54 GLPLFIHSIYSWIDL-----STRGYLWVWYIRNLIGSV 89
 DB 110 -----TMSRVAGEDRFGEKMKNTQLITFLIIGSIAGVLSINTI----- 152
 QY 90 HFFFLITLIVLITLYLWQTEGRKIMIRLHQIINEG---KDKMFLIEKLKL 141
 DB 153 -FPFLWVITLFLTW-----MSVFIKVPSVETNHGDQNDHSIKIKESLKI 201
 RESULT 3
 Y38A.MYGE STANDARD; PRT; 149 AA.
 ID Y38A.MYGE
 AC 092B71;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MG384.1.
 GN MG384.1.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 ON NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
 RA Nguyen D.T., Utterback T.R., Saudke D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403(1995).
 RN [2]
 RP IDENTIFICATION.
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
 RA Nguyen D.T., Utterback T.R., Saudke D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U39720; AAC71616.1; -
 DR TIGR: MG384.1; -
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 12 32
 FT POTENTIAL.
 FT POTENTIAL.

```

FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
SQ SEQUENCE 149 AA; 17731 MW; FOA7ABDD2562384 CMC64;

Query Match
Best Local Similarity 7.9%; Score 79.5; DB 1; Length 149;
Matches 36; Conservative 22; Mismatches 53; Indels 35; Gaps 7;

OY 11 WRASQMTFFILFFLP-SFGVCLTAITWRKPSADCGFRGLPFIHSIYSWIDL 69
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 15 WFLVPLLIYFLAFLEAFKTYI-----IWELIPECYSTAFLLIVLSGIIP----- 63
OY 70 STRPGYLWVWYIRNLIGSVHFFILTLITLITVL-----YQITGGRKIMIRLHE 122
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 64 -----MANNSW-----ILRFPLVILITLMSFLFNKMNIFLINSKAIITIA---E 108
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 123 QIINEGKDMFLIEKLIKQ-Q-DMERK 147
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 NLKTKGSKTKNRKQKIQIOSDLERK 134

RESULT 4
PNUC_SALTY STANDARD: PRT; 239 AA.
AC P24520;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein pnuC.
GN PNUC OR STM0757.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90330519; PubMed=2198247;
RA Foster J.W., Park Y.K., Penfound T., Fenger T., Spector M.P.;
RT "Regulation of NAD metabolism in Salmonella typhimurium: molecular
RT nadA-pnuC operon."
RT J. Bacteriol. 172:4187-4196(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Lt2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvanev E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT Lt2."
RN Nature 413:852-856(2001).
CC -1- FUNCTION: REQUIRED FOR NMN TRANSPORT ACROSS THE CYTOPLASMIC
CC MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- INDUCTION: REPPRESSED BY NADR.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 229.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M85180; -, NOT_ANNOTATED_CDS.
CC DR EMBL: AE008731; AAL19696.1; -.
CC DR PIR: D37753; D37753.
CC DR StyGene; SG10306; pnuC.

```

```

KW Transport; NAD; Transmembrane; Complete proteome.
FT TRANSMEM 22 42 POTENTIAL.
FT TRANSMEM 49 68 POTENTIAL.
FT TRANSMEM 72 89 POTENTIAL.
FT TRANSMEM 110 127 POTENTIAL.
FT TRANSMEM 158 177 POTENTIAL.
FT TRANSMEM 184 206 POTENTIAL.
FT CONFLICT 71 71 L -> R (IN REF. 1).
SQ SEQUENCE 239 AA; 27190 MW; D3A35FE6952AF86 CMC64;

Query Match
Best Local Similarity 7.9%; Score 79; DB 1; Length 239;
Matches 33; Conservative 26; Mismatches 51; Indels 24; Gaps 7;

OY 6 PPSKAMRASQMTFFILFFPSFGVCLTAITWR-----LKPSADCGFRGL 55
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 PRAMAMLAICVTAIGLMTRYIDPVAFITRAVAMQMGLOVTPVLPQ--DAPFWD 164
OY 56 PLFHSIYSWIDLSTRPGYL--WVWYIRNLIGSVH-----FILLITLITLITLYMQI 109
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 165 CMNVLSIYAMT--LMTKR-YVENMLMVIYINISVYIFALQGVAMSLEILITFI---A 218
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 110 TEGRKIMIRLHEQ 123
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 219 VNGSRLMINSARER 232

RESULT 5
SPAS_SHIFL STANDARD: PRT; 342 AA.
AC P40707;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Surface presentation of antigens protein spA5 (Spa40 protein).
GN SPAS OR SPA40.
OS Shigella flexneri, and
OS Shigella sonnei.
OC Plasmid 230 kb pWYSH6000.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Shigella.
OX NCBI_TaxID=623, 624;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=flexneri; STRAIN=serotype 2A;
RX MEDLINE=93224456; PubMed=8385666;
RA Sasakawa C., Komatsu K., Tohe T., Suzuki T., Yoshikawa M.;
RT "Eight genes in region 5 that form an operon are essential for
RT invasion of epithelial cells by Shigella flexneri 2a."
RT J. Bacteriol. 175:2334-2346(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S. sonnei; STRAIN=HM383;
RA Arakawa E., Kato J.I., Ito K.I., Watanabe H.;
RT Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: REQUIRED FOR SURFACE PRESENTATION OF INVASION PLASMID
CC ANTIGENS. COULD PLAY A ROLE IN PRESENTING THE TRANSLATION
CC COMPETENCE OF THE THREE IPA ANTIGENS. REQUIRED FOR INVASION AND FOR
CC SECRETION OF THE THREE IPA PROTEINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D13663; BAA02832.1; -.
CC DR EMBL: D50601; BAA09165.1; -.

```



```

SQ SEQUENCE 220 AA; 24604 MW; BA38B9644794B8C6 CRC64;
Query Match 7.6%; Score 76.5; DB 1; Length 220;
Best Local Similarity 23.3%; Pred. No. 4.9;
Matches 42; Conservative 34; Mismatches 59; Indels 45; Gaps 11;

OY 29 FTGVCTATIAITWRKPEADCGPRGLPFIHSIYSWIDLTSTRGCLYMWVIYNLIGS 88
   ||| : | | : | : | : ||| : | | : | | : | : | : | : | : |
Db 66 FPGLLVDVKVTSIR-QPOSSC-PEPSAROKIFGL-----GYSLLFVVYDKDINS 112
   ||| : | | : | : | : ||| : | | : | | : | : | : | : | : |
OY 89 VHFPIILLIYLITYLQWOTTEGRKIMIRILHEQIINEG-KDKM--PLIEKLTIADME 145
   || | | : | : | : | : | : ||| | : | : | : | : | : | : | : |
Db 113 TN--RTATNLNLIHTIYSAERTADFO-WTRGIIRNLANEGKNKDILAMSDRNLPYDETE 169
   ||| : | | : | : | : ||| : | | : | | : | : | : | : | : |
OY 146 KKANPSSLVERREVEQGFLHL-----GEHDSIDLRSRVSVOEGNP 188
   ||| : | | : | : | : ||| : | | : | | : | : | : | : | : |
Db 170 -AGNVAAITLL--RNDPMOGFLTISNALQWRLQYGVIERAGEEDGIL-----TYRRNP 220
   ||| : | | : | : | : ||| : | | : | | : | : | : | : | : |

RESULT 8
YEDO_ECOLI STANDARD; PRT; 564 AA.
AC P76330; P94746; MEDLINE=97251358; PubMed=9097040;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yedQ.
GN YEDO OR B1956.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9276503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew D.J.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RP [2]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Saba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasubraman S., Tagami H., Takekida J., Takemoto K., Wada C.,
RA Yamamoto Y., Horikuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 Genome
Corresponding to the 4.0 x 10^6 bp Region on the Linkage Map.";
RL DNA Res. 3:379-392(1996).
CC -1 SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1 SIMILARITY: BELONGS TO THE YAIC / YEIN (E.COLI), YHCK (B.SUBTILIS)
FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; AE000287; AAC75022.1; ALT_INIT.
DR EMBL; D90835; BAA15784.1; -.
DR EcoGene; EG14040; yedo.
DR InterPro; IPR000160; GGDEF.
DR Pfam; PF00990; GGDEF_1.
DR SMART; SM00267; DUFL_1.

```

```

DR  TIGRFAMs:TTGR00254: GGDFF: 1.
KM  Hypothetical protein; Transmembrane; Complete proteome.
FT  TRANSMEM 20 40 POTENTIAL.
FT  TRANSEM 360 380 POTENTIAL.
SQ  SEQUENCE 564 AA; 64283 MW; 05FB02C1BE2A8938 CRC64;

Query Match 7.6%; Score 76; DB 1; Length 564;
Best Local Similarity 26.0%; Pred. No. 15;
Matches 38; Conservative 25; Mismatches 43; Indels 40; Gaps 8;

QY 64 SWITFLNTR--PGYLMVWVIRNLIIGSVHFFILITLYLITTYLYWQ--TTEGRKIMIR 118
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 11 SWLKRLARRLRGPHV-----VNLCFI--VLLFSTLLFWREVVLEDAVYISSQ 56
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 119 LLHQIINEGDKR--MFLIEKLIKIDMEKRA-----NPSSLVIERREVEQGFILHGEH 171
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 57 RNHLEENVANALDKHLQYVNVKLIPLFRNGMRALVAPLDFTSLRDVATFEQ----HRDEH 112
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

QY 172 DGSIDLRSRS-----VQSEN 187
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 113 AMKIETLNRRRTLPVNGVSDALVSEGN 138
    ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 9
ID ALA8_ARATH STANDARD; PRT; 1189 AA.
AC 09LK90;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Potential phospholipid-transporting ATPase 8 (EC 3.6.3.1)
DE (Aminophospholipid flippase 8).
GN ALA8 OR A13G27870 OR K16N12.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eusterosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneke T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
RT and BAC clones."
RL DNA Res. 7:217-221(2000).
CC -1- FUNCTION: INVOLVED IN TRANSPORT OF PHOSPHOLIPIDS (POTENTIAL).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
CC (EI-E2 ATPASES). SUBFAMILY IV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP000371; BAB02533.1; -
CC InterPro: IPR0014757; ATPase_EI-E2.
CC InterPro: IPR001454; H1gnase/Hydr1ase.
CC Pfam; PF00702; Hydro1ase; 1.
CC PRINTS; PR00119; CATATPASE.
CC PROSITE; PS00154; ATPASE_EI_E2; 1.
CC Hydro1ase; Transmembrane; Phosphorylation; ATP-binding; Metal-binding;
CC Magnesium; Multigene family.
CC DOMAIN 1 71 CYTOPLASMIC (POTENTIAL).
CC TRANSEM 72 93 POTENTIAL.
CC DOMAIN 94 97 EXTRACELLULAR (POTENTIAL).
CC TRANSEM 98 120 POTENTIAL.

```

```

FT DOMAIN 121 299 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 300 321 POTENTIAL.
FT DOMAIN 322 358 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 359 376 POTENTIAL.
FT DOMAIN 377 920 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 921 940 POTENTIAL.
FT DOMAIN 941 954 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 955 974 POTENTIAL.
FT DOMAIN 975 1004 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1005 1027 POTENTIAL.
FT DOMAIN 1028 1040 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1041 1063 POTENTIAL.
FT DOMAIN 1064 1069 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1070 1090 POTENTIAL.
FT DOMAIN 1091 1107 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1108 1132 POTENTIAL.
FT DOMAIN 1133 1189 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 424 424 PHOSPHORYLATION (BY SIMILARITY).
FT METAL 865 865 MAGNESIUM (BY SIMILARITY).
FT METAL 869 869 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 1189 AA; 135309 MW; DCF951CF3FC5EB3 CRC64;

Query Match 7.5%; Score 75.5; DB 1; Length 1189;
Best Local Similarity 23.4%; Pred. No. 40;
Matches 44; Conservative 30; Mismatches 71; Indels 43; Gaps 9;

QY 6 PSKAMRASQMTFFFLFPSPFTGLTATL-----TWRLKPSADCGPRGLPLFT 59
DB 285 PSKRSKIEKKMQDIITYLE-----SLIVIAFTGSVFGLATRRDMSGKLR----- 333
QY 60 HSLSYWDITSTRPGYIVWVYIRNLIGSVHFFILTLV---LIITYLYKQITTEGRKIM 116
DB 334 ----RW----LRPDHTTYEDPRRAVAAAFHFLTALMGYLIPISLVSTIE- 380
QY 117 IRLHEQINEGKDKMLKLIKLODEKKAAPSSIVLRRERQGGFLHEDGSILD 176
DB 381 VKVLQSLFTIQDDE-WYHEE-----IDPRKARTSNINELGOVD-----TILSDKGTGLT 430
QY 177 LRSRSVQ 184
DB 431 CNSMEFYK 438

RESULT 10
MOK1_SCHPO STANDARD: PRT: 2410 AA.
AC Q9USRB; Q9URTS; Q9URK2;
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DE Cell wall alpha-1,3-glucan synthase moki (EC 2.4.1.183).
GN MOK1 OR AGS1 OR SPC338.01C OR SPC17A7.01.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA Katayama S., Dai H., Arellano M., Perez P., Toda T.;
RT "Fission yeast alpha-glucan synthase MOK1 localizes closely with actin
RT and play a role essential for cell morphogenesis and protein kinase C
RT function."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Godle A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

```

```

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Nibbel D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabbittowitch E.,
RA Rutherford K., Ruter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Filicz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wandubert R., Purnelle B.,
RA Corfau A., Cadieu E., Diano S., Gloux S., Lelaure V., Kottler S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -1- CATALYTIC ACTIVITY: UDP-glucose + {alpha-D-glucosyl-(1,3)}(N+1) =
CC UDP + {alpha-D-glucosyl-(1,3)}(N+1)
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AB019183; BAA34054.1; -
DR EMBL: AL023781; CAI19332.1; -
DR EMBL: AL049472; CAB39330.1; -
DR InterPro: IPR000461; Alpha_1,3-amyase.
DR InterPro: IPR001296; Glycosyl_transf_1.
DR Pfam: PF00128; alpha-amyase; 1.
DR Pfam: PF00534; Glycosyl_transf_1; 1.
DR Cell wall; Transferase; Glycosyltransferase.
FT CONFLICT 256 256 F->Y (TN REF. 2)
SQ SEQUENCE 2410 AA; 272120 MW; 8BD944BAE3A9A5C5 CRC64;

Query Match 7.5%; Score 75.5; DB 1; Length 2410;
Best Local Similarity 24.3%; Pred. No. 87;
Matches 33; Conservative 13; Mismatches 49; Indels 41; Gaps 6;

QY 7 PSKAMRASQMTFFFLFP-----PPSFTGLTATITWRLKPSADCGPR 53
DB 2153 PSITWPLSAVSILIFALLFGLEPRYRQLSGISIPARYKSLRRKLVW----- 2200
QY 54 GULFLHSIY---SWDITSTR-GYLWVWYIRNLIGSVHFFILTLVLIITYLYKQI 109
DB 2201 ----FCISYFLQNFWSLSLNGRMSYSLMD-----IGNHMQIFLLIVAFYIVL-WAL 2248
QY 110 TEGRKIMIRLHEQII 125
DB 2249 LIGVLAMISRTSHWII 2264

RESULT 11
PANI_HUMAN STANDARD: PRT: 422 AA.
AC Q96RD7; Q75968; Q96RS5; Q96L77; Q96AM;
DT 15-JUN-2002 (Rel. 41; Created)
DT 15-JUN-2002 (Rel. 41; Last sequence update)
DE Pannekin 1.
DE PANXI OR MRS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```



```

FT REPEAT 335 358 LRR 10.
FT REPEAT 359 384 LRR 11.
FT REPEAT 386 411 LRR 12.
FT REPEAT 412 436 LRR 13.
FT REPEAT 438 456 LRR 14.
FT REPEAT 457 476 LRR 15.
FT REPEAT 477 499 LRR 16.
FT REPEAT 501 521 LRR 17.
FT REPEAT 523 546 LRR 18.
FT DOMAIN 639 784 TIR.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 784 AA; 90204 MW; 59FDB9DFA7984C18 CRC64;

Query Match 7.5%; Score 75; DB 1; Length 784;
Best Local Similarity 23.0%; Pcore No. 28;
Matches 45; Conservative 23; Mismatches 58; Indels 70; Gaps 11;

QY 21 IFLFFPSFTGVCTLAITW-----RLKPSADGPPRGDPLFIHSIYS----- 64
   ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 598 LFLLL--LTGVCHRFHGLMYKMMWMLQAKRKPR--APRRDICYAFVSERDSY 653
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 65 WIDTLESTR-----PGYLVWVWYIRNLIGSV---HFFILITLVLII 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 654 WVENLWVQLEHNPPEKCLHAKRDPIG---KWIIDNIDISIEKSH---KTIPEVSE 705
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 TYVYWOITEGRKIMIRLHQIINEGDKMKLEIKLQDMKKNPSSLYERR--- 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 706 NEVK---SECKKEIPEFSHFRLENDENNAI---LILPEIDKKAIPQFCALKRKIMNT 758
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 159 -----EVEQGF 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 759 KYLEWPEVDETQOGEF 774
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
TLR2_HUMAN STANDARD; PRT; 784 AA.
ID TLR2_HUMAN
AC 060603; O15454;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toll-1-like receptor 2 precursor (Toll/interleukin 1 receptor-1like protein 4).
DE TLR2 OR TIR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte, and Prostate;
RA MEDLINE=9826144; PubMed=9596645;
RA Chaudhary F.M., Ferguson C., Nguyen V., Nguyen O., Massa H.F., Eby M.,
RA Jasmin A., Trask B.J., Hood L., Nelson P.S.;
RT "Cloning and characterization of two Toll/interleukin-1 receptor-1like
RT genes TIR3 and TIR4: evidence for a multi-gene receptor family in
RT humans";
RL Blood 91:4020-4027(1998).
RN (2)
RP SEQUENCE FROM N.A.
RA MEDLINE=98118556; PubMed=9435236;
RA Rock F.L., Hardiman G., Timans J.C., Kastelein R.A., Bazan J.F.;
RT "A family of human receptors structurally related to Drosophila
RT Toll-1";
RL Proc. Natl. Acad. Sci. U.S.A. 95:588-593(1998).
RN (3)
RP SEQUENCE FROM N.A., AND RESPONSE TO LIPOPOLYSACCHARIDE.
RC TISSUE=Fetal Lung;
RA MEDLINE=98421677; PubMed=9751057;
RA Yang R.-B., Mark M.R., Gray A., Huang A., Xie M.H., Zhang M.,
RA Goddard A., Wood W.L., Gurney A.L., Godowski P.J.;

```

```

RT "Toll-1-like receptor-2 mediates lipopolysaccharide-induced cellular
RT signalling.";
RL Nature 395:284-288(1998).
RN (4)
RP RESPONSE TO BACTERIAL LIPOPROTEINS.
RA MEDLINE=99357867; PubMed=10426996;
RA Aliprantis A.O., Yang R.-B., Mark M.R., Suggest S., Devaux B.,
RA Radolf J.D., Kimpel G.R., Godowski P.J., Zychlinsky A.;
RT "Cell activation and apoptosis by bacterial lipoproteins through
RT Toll-1-like receptor-2.";
RL Science 285:736-739(1999).
RN (5)
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF TIR DOMAIN, AND MUTAGENESIS.
RA MEDLINE=20531768; PubMed=11081518;
RA Xu Y., Tao X., Shen B., Horng T., Medzhitov R., Manley J.L., Tong L.;
RT "Structural basis for signal transduction by the Toll/interleukin-1
RT receptor domains.";
RL Nature 408:111-115(2000).
CC -1- FUNCTION: Cooperates with MD-2 to mediate the innate immune
CC response to bacterial lipoproteins and other microbial cell wall
CC components. Acts via MyD88 and TRAF6, leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response. May
CC also promote apoptosis in response to lipoproteins. Recognizes
CC mycoplasma macrophage-activating lipopeptide-2kD (MALP-2), a
CC soluble tuberculosis factor (STF), phenol-soluble modulin (PSM)
CC and B.burfordi outer surface protein A lipoprotein (OspA-L)
CC cooperatively with TLR6.
CC -1- SUBUNIT: Binds MD-2 and TLR6 via the extracellular domain. Binds
CC MyD88 via their respective TIR domains.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in peripheral blood
CC leukocytes, in particular in monocytes, in bone marrow, lymph node
CC and in spleen. Also detected in lung and in fetal liver. Levels
CC are low in other tissues.
CC -1- PTM: N-glycosylated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -1- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -1- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: AF051152; AAC34377.1; -.
DR EMBL: U88878; AAC34133.1; -.
DR Genew; HGNC:11848; TLR2.
DR MIM: 603028; -.
DR PDB: 1FYW; 22-NOV-00.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR003581; LRR_Typ.
DR InterPro: IPR000157; TIR_domain.
DR Pfam: PF00560; LRR; 6.
DR Pfam: PF01463; LRCT; 1.
DR Pfam: PF01582; TIR; 1.
DR SMART: SM00082; LRCT; 1.
DR SMART: SM00369; LRR_Typ; 2.
DR SMART: SM00255; TIR; 1.
DR PROSITE: PS0104; TIR; 1.
KW Receptor; Immune response; Inflammatory response; Signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein;
KW 3D-structure.
KW SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 784 TOLL-LIKE RECEPTOR 2.
FT DOMAIN 19 588 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 589 609 POTENTIAL.
FT DOMAIN 610 784 CYTOPLASMIC (POTENTIAL).
FT REPEAT 51 74 LRR 1.
FT REPEAT 75 98 LRR 2.

```

```

FT REPEAT 99 122 LRR 3.
FT REPEAT 124 147 LRR 4.
FT REPEAT 148 172 LRR 5.
FT REPEAT 173 196 LRR 6.
FT REPEAT 221 244 LRR 7.
FT REPEAT 359 384 LRR 8.
FT REPEAT 412 436 LRR 9.
FT REPEAT 438 456 LRR 10.
FT REPEAT 457 477 LRR 11.
FT REPEAT 477 499 LRR 12.
FT REPEAT 501 521 LRR 13.
FT REPEAT 521 544 LRR 14.
FT CARBOHYD 639 784 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 114 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MOTAGEN 681 681 P->F: ABOLISHES THE INTERACTION WITH MYD88. NO EFFECT ON OLIGOMERIZATION OR ON THE STRUCTURE OF THE TIR DOMAIN.
FT CONFLICT 726 726 D -> E (IN REF. 2).
SQ SEQUENCE 784 AA; 89837 MM; 7DBEB24CF1F8B CRC64;

Query Match 7.5%; Score 75; DB 1; Length 784;
Best Local Similarity 23.0%; Pred. No. 28;
Matches 45; Conservative 23; Mismatches 58; Indels 70; Gaps 11;

QY 21 FFLFFPSFTGVLCITLITW-----RLKPSADCGPFGPLFIHSITS----- 64
DB 598 LFLIL--LGLVLCRHREHGLWMYMKMMALQAKRPK--APSRNICDAFVSYSERDAY 653
QY 65 WIDFLSTR-----PGYLWVYIYRNIGSV---HFFILTLIVLII 102
DB 654 WVENLMQOLENPNPRLCLHKRDFIPG---KWIIDNIIDSLEKSH---KIVFVISE 705
QY 103 TYLWQITEGRKIMIRLLHQIINEGDKMFLIEKLQDMKRNANSSLVLERR--- 158
DB 706 NFVK---SEMKYEIDFSHFLRFDENNDAI---LILLEDIEKKAIQRCCKLRKIMNT 758
QY 159 -----EVEQOGF 165
DB 759 KTYLEWPMDEAQRGEF 774

RESULT 14
DSBL_BACSU
ID DSBL_BACSU STANDARD: PRT; 148 AA.
AC 064037; 031985;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative protein-disulfide oxidoreductase from Spheta2 prophage.
GN YOLK.
OS Bacillus subtilis, and
OS Bacteriophage SPBc2.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423, 66797;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=B.subtilis; STRAIN=168;
RX MEDLINE=9824026; PubMed=9579063;
RA Regamey A., Karamata D.;
RT "The N-acetylmuramoyl-L-alanine amidase encoded by the Bacillus subtilis 168 prophage SP beta.";
RL Microbiology 144:885-893(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=B.subtilis; STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Allont G.,
RA Azavedo V., Bertoletti M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

```

```

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devigne K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleton N.,
RA Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Meuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Potworlik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadele Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scollone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terstra P., Tognoni A.,
RA Trosato V., Uchiyama S., Vandenhof M., Vannier F., Vassartoli A.,
RA Viari A., Wanduit R., Wedler E., Wedler H., Weitenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus subtilis.";
RT Nature 390:249-256(1997).
RL [3]
RN SEQUENCE FROM N.A.
RP SPECIES=Phage SPBc2;
RA Lazarevic V., Duesterhoeft A., Soldo B., Hilbert H., Meuel C.,
RA Karamata D.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: REQUIRED FOR DISULFIDE BOND FORMATION IN SOME PROTEINS (POTENTIAL).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -I- SIMILARITY: BELONGS TO THE DSBB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF021803; AAC38303.1; -.
CC DR EMBL; 299115; CAB14062.1; -.
CC DR EMBL; AF020713; AAC12996.1; -.
CC DR Subtilist; BG13587; YOLK.
CC DR InterPro: IPR003752; DSBB.
CC DR Pfam; PF02600; DSBB; 1.
CC KW Hypothetical protein; Oxidoreductase; Redox-active center;
CC FT Electron transport; Chaperone; Transmembrane.
CC FT TRANSMEM 8
CC FT TRANSMEM 28
CC FT TRANSMEM 65
CC FT TRANSMEM 85
CC FT TRANSMEM 110
CC FT TRANSMEM 130
CC FT DISULFID 36
CC FT DISULFID 39
CC FT REDOX ACTIVE (BY SIMILARITY).
SQ SEQUENCE 148 AA; 17139 MM; 25676A3B9C9AF60C CRC64;

Query Match 7.4%; Score 74.5; DB 1; Length 148;
Best Local Similarity 22.1%; Pred. No. 4.8;
Matches 23; Conservative 22; Mismatches 40; Indels 19; Gaps 4;

QY 20 FFLFFPSFTGVLCITLIT--IWRLKPSADCGPFGPLFIHSIYWIDTLSTRP-GYLV 77
DB 9 FFLFLFLSFSGTMASTLYSEIMHPKPCVLC-----WQRIFLXPPIIL 53
QY 78 VVWYIRNLIGSVHFFILTLIVLITYLW--QITEGKIMIRL 119
DB 54 LIGLKKDNLNIFVYVFLSIGLIIAFYHIIDLOSKSVCEI 97

```

```

RESULT 15
Y464_MYCGE          STANDARD;          PRT;          385 AA.
ID   Y464_MYCGE
AC   P47702;
DT   01-OCT-1996 (rel. 34, Created)
DT   01-OCT-1996 (rel. 34, Last sequence update)
DT   15-JUN-2002 (rel. 41, Last annotation update)
DE   Hypothetical protein MG464.
GN   MG464.
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RX   MEDLINE=96026346; PubMed=7569993;
RA   Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA   Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA   Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA   Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA   Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA   Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT   "The minimal gene complement of Mycoplasma genitalium.";
RL   Science 270:397-403(1995).
CC   -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC   -1- SIMILARITY: To A.M.CAPRICOLUM HOMOLOG.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@sib-sib.ch).
CC   -----
DR   EMBL: U39728; AAC72484.1; -
DR   TIGR: MG464; -
DR   InterPro: IPR001708; 60kDa_innermemb.
DR   Pfam: PF02096; 60KD_TMP; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM          39      59      POTENTIAL.
FT   TRANSMEM          150     170     POTENTIAL.
FT   TRANSMEM          222     242     POTENTIAL.
FT   TRANSMEM          270     290     POTENTIAL.
FT   TRANSMEM          325     345     POTENTIAL.
SQ   SEQUENCE          385 AA; 44207 MW; BCB2964F7842893A CRC64;

Query Match          7.4%; Score 74.5; DB 1; Length 385;
Best Local Similarity 28.3%; Pred.No.14;
Matches 41; Conservative 19; Mismatches 28; Indels 57; Gaps 9;

OY   56 PLFIHSIYSWIDLTSTRPGIWWVW-----IY--RNLIQS--VHFFILTLIVLI 101
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB   103 PYTFPSDY-----TLAYGPFYGFVWPAQAQIVLPIMATRPVPLSGVGLGFNMTLSLIVLL 158
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY   102 IYLYWQITEGRKIMRLIHLHQIINEGKDKMFLIEKLKIQDMEKKRANPSSLVLEEREVE 161
    : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB   159 -----LIVR-----LITIVITLN-----STLLEKMN-E 181
    : : : | : : : | : : : | : : : | : : : | : : : | : : : |
OY   162 QOGFLH--LGEHDSGLDLRSRSVQ 184
    | | | | | : : : | : : : | : : : | : : : | : : : | : : : |
DB   182 VOGKLAELNKKYKGAIDLQSKRRNQ 206
    | | | | | : : : | : : : | : : : | : : : | : : : | : : : |

```

Search completed: November 9, 2002, 04:31:25
 Job time : 42 secs